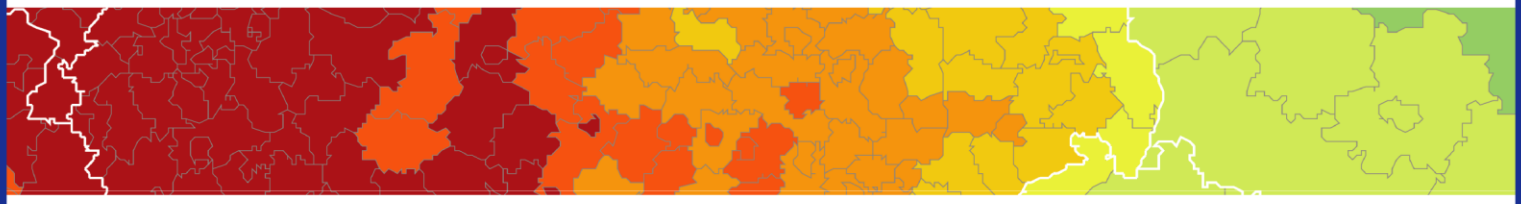
ESPON

EUROPEAN UNION

Co-financed by the European Regional Development Fund

Inspire policy making by territorial evidence



Potentials of big data for integrated territorial policy development in the European growth corridors (Big Data & EGC)

Targeted Analysis

Synthesis Report

28/06/2019

This Targeted analysis activity is conducted within the framework of the ESPON 2020 Cooperation Programme, partly financed by the European Regional Development Fund.

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.

This delivery does not necessarily reflect the opinion of the members of the ESPON 2020 Monitoring Committee.

Authors
Helka Kalliomäki, Centre for Collaborative Research , University of Turku (Finland)
Ira Ahokas, Nicolas Balcom Raleigh, Finland Futures Research Centre, University of Turku (Finland)
Jukka Heikkonen, Pekko Lindblom, Paavo Nevalainen, Ville Keränen, The Department of Information Technology, University of Turku (Finland)
Siiri Silm, Anto Aasa, Mobility Lab, University of Tartu (Estonia)

Advisory Group
Nicolas Rossignol, ESPON EGTC (Luxembourg)
Antti Vasanen, Regional Council of Southwest Finland (Finland)
Dino Keljalic, Region Örebro (Sweden)
Liis Vahter, Ministry of Economic Affairs and Communications (Estonia)

Information on ESPON and its projects can be found on www.espon.eu.

The web site provides the possibility to download and examine the most recent documents produced by finalised and ongoing ESPON projects.

This delivery exists only in an electronic version.

Contact: info@espon.eu

Disclaimer:
This document is a synthesis report.

The information contained herein is subject to change and does not commit the ESPON EGTC and the countries participating in the ESPON 2020 Cooperation Programme.

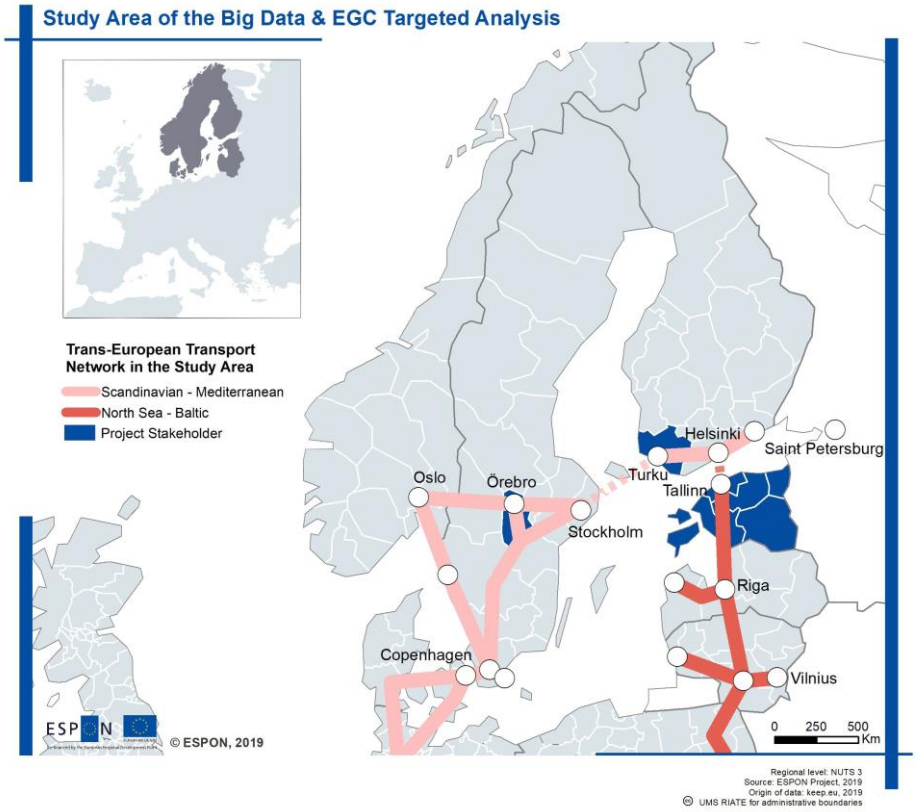
The final version of the report will be published as soon as approved.

Need for a broader understanding functional growth corridors

The potentials of big data for corridor-based development are evident and widely discussed, yet there is a need to develop specific approaches to utilising big data to inform growth corridor policy. As a setting, the functions of corridors are influenced by multiple levels of governance and varying scales involving many kinds of flows and interactions. These overlapping and interlinked flows and interactions cannot be fully captured by using traditional data sources. Yet, big data and new data sources could be used to produce new insights and evolve a more comprehensive understanding of corridors. Such a broader understanding of flows and interactions is needed to support comprehensive corridor governance. Yet, there are currently very few detailed analyses about the functionality of growth corridors. New data sources can build new forms of evidence for corridor-based development.

The aim of the Big Data & EGC Targeted Analysis was to strengthen the knowledge-base for how to approach evidence-based planning in European growth corridors that often build on the Ten-T transport infrastructure network. This analysis covered the so-called Northern Growth Zone (NGZ) that stretches from Oslo via Örebro to Stockholm; Turku to St. Petersburg; and following the northern parts of Scandinavian-Mediterranean (ScanMed) TEN-T Corridor (Map 1).

Map 1: Study area of the targeted analysis.



By examining potentials for big data to support corridor-related policy-making in research area covering Finland, Sweden and Estonia, the project answered to the needs of project

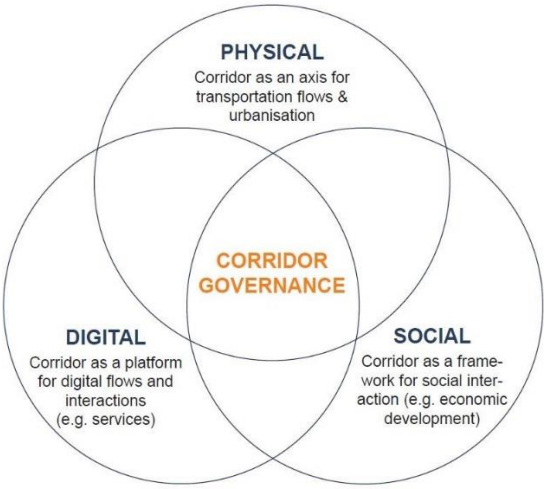
stakeholders of the Regional Council of Southwest Finland, Region Örebro in Sweden, and the Estonian Ministry of Economic Affairs and Communications. According to these stakeholders, the key policy dimensions related to corridor development that would benefit from big data are: 1) infrastructure and connectivity planning; 2) regional economic development, and; 3) land-use planning. Using new datasets in policy-making could fill the gaps in the evidence-base and connect policy silos. Understanding the motivations and drivers of, e.g., physical mobility highlights the significance of more comprehensive examination of spatial connectivities.

This analysis sought to generate new knowledge about how big data can be used to expose functionalities of the NGZ. A goal of the analysis was to demonstrate and test new methods for producing insights from unusual datasets across three case studies and a hackathon. This analysis contributes to larger efforts to develop new ways for supporting policy-making with timely and highly detailed evidence.

A framework for comprehensive corridor governance

Gaining a more comprehensive view of corridor flows and interactions can be achieved by exploring an issue from multiple angles. To this end, a conceptual model was produced in the targeted analysis to help policymakers consider what kinds of data could be used to support corridor governance based on evidence (Figure 1). The framework describes three overlapping functional dimensions of corridor development:

Figure 1: Conceptual tool for broadening the perspective of corridor functionalities



1) the **physical** dimension emphasizing transportation flows and urbanisation;
2) the **digital** dimension realizing digital flows and interactions as an active part of the corridor functions;
3) and, the **social** dimension viewing the corridor as a framework for social interaction.

This conceptual framework was applied to produce a categorisation of key flows and interactions which are highly relevant to corridor governance and policymaking. This categorization presented in the Big Data & EGC Targeted Analysis includes examples of what

kinds of datasets are available concerning each kind of flow and interaction. The main purpose of this categorisation is to serve as a starting point for considering what kinds of existing or new data to include when designing data-driven inquiries to support corridor-wide policymaking.

Case studies examining the potentials of new data sources

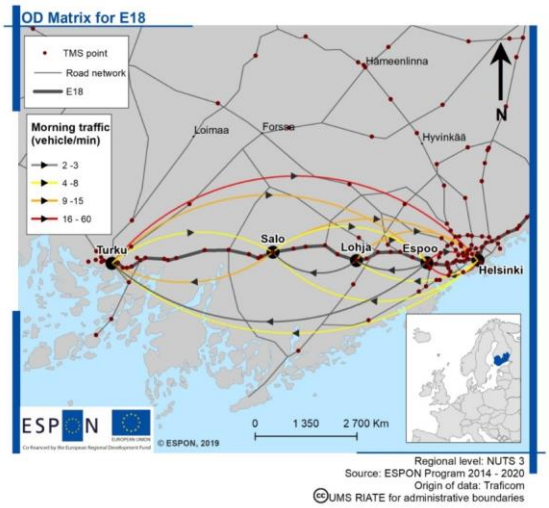
This targeted analysis included four case studies – one using data collected from induction loop sensors in a highway to produce O-D matrixes and a predictive model; another producing a spatial network analysis of project collaborations; a third one processing passive mobile positioning data to produce a database of daily O-D matrices visualised on road systems; and finally a hackathon that invited students to explore policy relevant uses of big data. The Synthesis Report presents main findings of the three case studies.

Traffic measurement data provides an important aspect for wider flow analytics

The objective of this case study was to explore the potentials of automatically collected traffic data for producing a predictive model of traffic flows on E18 in Finland and based on it, origin-destination matrix for the highway. The materials of the case study consisted of automated traffic intensity measurement (ATM) data provided by the Finnish Transport and Communications Agency (Traficom). Similar forms of data is passively and automatically gathered, e.g., by Swedish Transport Administration (Trafikverket) and Estonian Road Administration (Maanteeamet) which indicated cross-cutting value of developing an analytical tool across the whole NGZ.

In the case study, a new model was developed to indirectly estimate the Origin-Source (O-D) matrix, which has not been possible earlier without direct observation methods like traffic polls and a full coverage of road junctions by either traffic sensors or human observers (Map 2).

Map 2: Traffic flows between Turku and Helsinki during a work week in 2017.



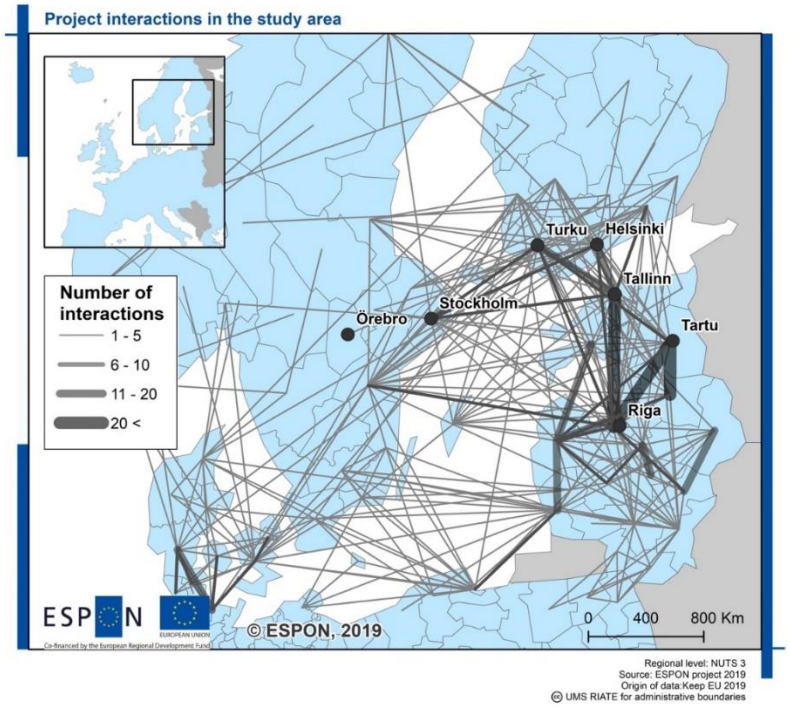
Traffic flows at 7:30-9:30 on Mon - Thu Jan 23rd - 26th 2017 between Turku and Helsinki along E18.

An O-D matrix can support policy-makers in analysing dependencies, similarities, synchronizations and differences in flows between cities and communities. Because the model is geographically generic, it could be used anywhere where similar traffic data are available. Now that it is built, it could be combined with data depicting wider corridor dynamics such as key economic variable. In summary, the outcome of the study in the form of O-D Matrixes and the predictive model can increase understanding of how traffic flows in growth corridors relate to other functions over time.

Social network analysis reveals a different side of functionality

In the second case study ERDF project partnership data was analysed to understand the collaboration dynamics in the NGZ. The data was downloaded from the keep.eu portal, which is a database providing public access to details about projects funded by ERDF programs including the partners involved, project budgets, thematic categories, and locations. The analysed dataset covered 2,353 projects conducted by 18,318 partners across the EU. The visualisation of the results clearly show the spatial imbalance of collaboration activity in the study area of the targeted analysis (Map 3).

Map 3: Project interactions in the study area of the Big Data & EGC Targeted Analysis



The approach used to produce these insights from the data could be used for larger scale trend analysis concerning project-related collaborations within European growth corridors (e.g. TEN-T). In the NGZ, results from the conducted analysis can be used to support, e.g., the development of funding instruments aimed at promoting interregional project collaborations especially within, and with, the western parts of the corridor.

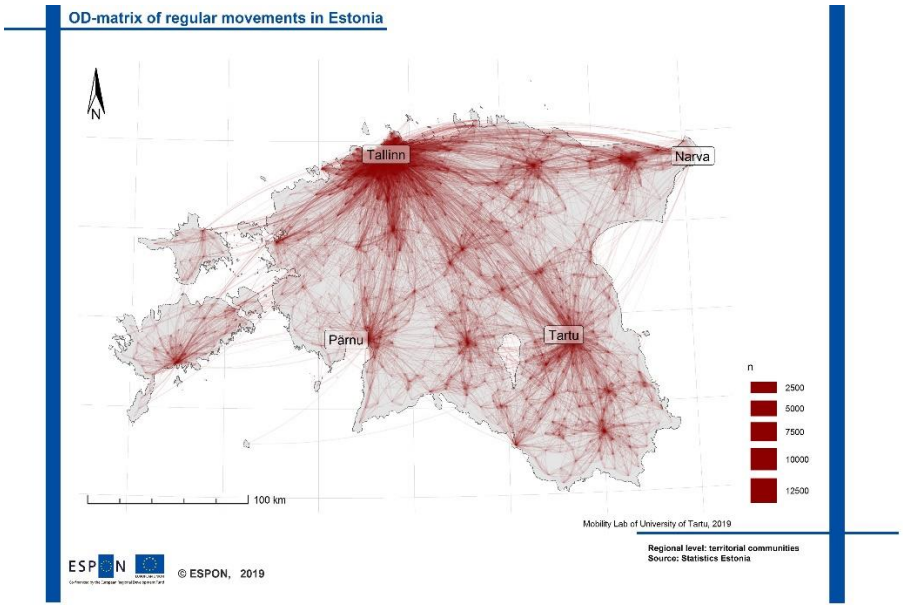
Altogether, an increased understanding of collaboration dynamics as a form of social interaction can help policymakers formulate policy-measures to support the strategic development of

functional areas. Furthermore, a wider understanding of these kinds of interconnections among regions opens possibilities to increase the legitimacy of, and commitment to, transnational corridor development. The experiments conducted in this case study demonstrate how big data analysis can reveal the spatial dimension of partnership networks which can support integrated policy-making that fosters corridor cohesion. Similar forms of social network analysis using other kinds of datasets - such as consortium, social media, or board member data - also could provide interesting insights on territorial development.

Mobile positioning data reveals the dynamics of mobility within spatial structures

The objective of the third case study was to develop a methodology for producing an everyday mobility database which contains OD-matrices of movements between territorial communities of Estonia mapped to road networks (see Map 4). This was done in collaboration with the Ministry of Economic Affairs and Communications. Before this case, mobile positioning data had already proven to be a useful data source for studying mobility related facets of the society and/or smaller groups and the research team leading this part of our analysis has many years of experience working with it.

Map 4: Regular movements in Estonia according to the mobile positioning data.



The data has been produced for many years and has a fine temporal resolution, meaning it includes observations at small time intervals. This makes it suitable for revealing short-term or seasonal differences in the numbers of people moving among their key daily locations such as work and home. In addition, this kind of dataset can include user-provided demographic details such as gender, approximate age, nationality etc. which could be analysed to gain insights into how mobility patterns differ for, e.g., tourists, young people, etc. Furthermore, the possibilities of utilising mobile positioning data for cross-border settings of the study area are becoming more possible, as this kind of data is increasingly available in neighbouring nations of Finland, Estonia and Sweden.

This case functions as a best practice example for Europe. The kind of dataset it relies upon and the analytical approach could be widely applied across the EU. This kind of data exists in all European nations and the process developed in this case study for analysing it is replicable. For example, the kind of dataset and analytical approach would be suitable for analysing mobility patterns in whole TEN-T corridors. The case study emphasizes how using mobile positioning data can reveal connectivities that cannot be captured by using traditional statistical datasets. This could significantly improve the understanding of the mobility dynamics of growth corridors. Therefore, the Estonian experiences related to the utilisation of mobile positioning data in national transportation planning have to be further studied to identify future potentials, best practices, and challenges related to utilising this dataset.

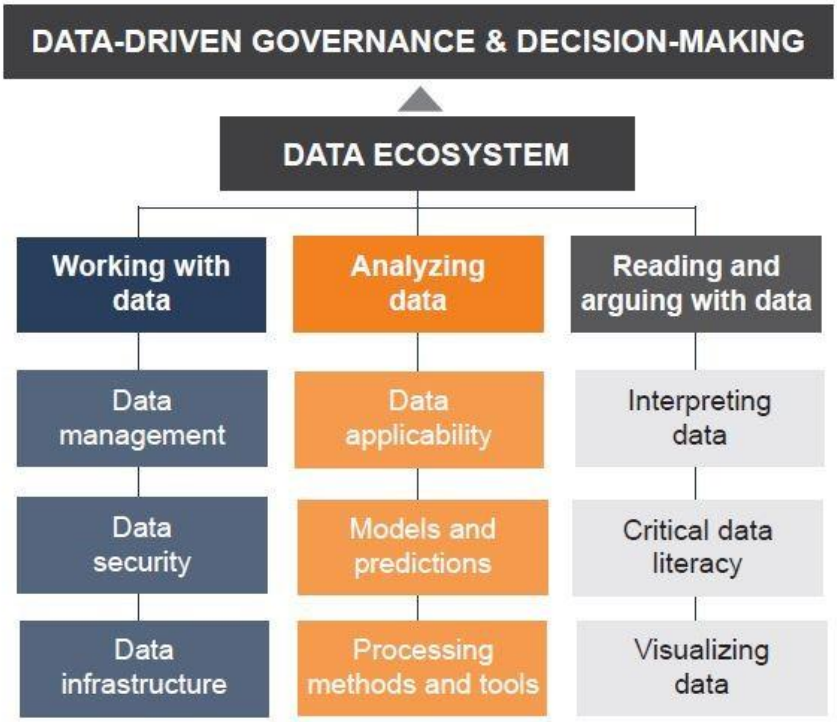
From data management to data-driven governance

In order to drive the use of big data in governance and decision-making across a corridor, it is important to aim capacity building practices at many levels – from singular organisations to corridor wide networks. Developing a data ecosystem for data-driven corridor governance requires building different data literacy capacity areas (see Figure 1). Overall, becoming data driven requires a mental shift from data management to data-driven governance. When organisational cultures become highly competent in using big data and aware of how data-derived insights are produced, they will be able to broaden the evidence-base of territorial development. Alongside this capacity building, it is highly important to incorporate ethical guidelines and best practices in the workplace so to help avoid making mistakes that can harm people (e.g. taking action based on a biased algorithm). An organisation's ability to use big data for evidence-based policy-making in corridors cannot rely on only a few technically skilled employees. While having teams of high competence is helpful, all employees and stakeholders of an organisation need awareness of how to use big data so any big data insights produced by any other actor can be appropriately scrutinised before actions are taken based upon it. Experts who are capable of transforming raw data into appropriable data and action are helpful to all parts of an organisation.

To establish data-driven governance in growth corridors, promoting data literacy at all levels of organisations and to the wider public is key. These data literacy skills should be emphasized:

- **working with data**: people need skills in working with data, including how to locate high quality datasets, how to read meta-data, and how essential functions work in datasets such as joins, sorts and queries.
- **analysing data**: before any data analysis process, many steps are required, beginning with selecting a dataset and which of its parts to analyse. To analyse a dataset, many choices need to be made regarding the approach to use, fields to include, etc. Even if analysing data is not a person's daily job, knowing how it is done supports participation in discussions about data and designs of data-driven inquiries.
- **reading and arguing with data**: A widespread ability to read and argue with data is essential to fostering a work culture that can function based on data. Arguing with data refers to the ability to visualize the results of analytics in an understandable way for policy-makers.

Figure 1. Recommendations for capacity building.



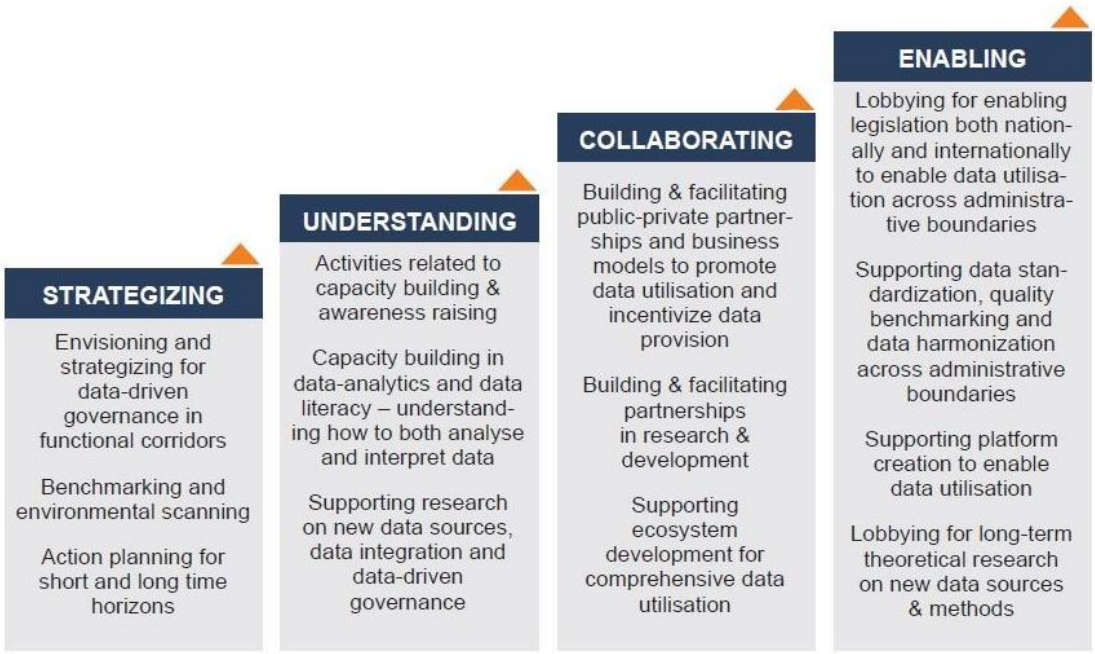
Summary of policy recommendations

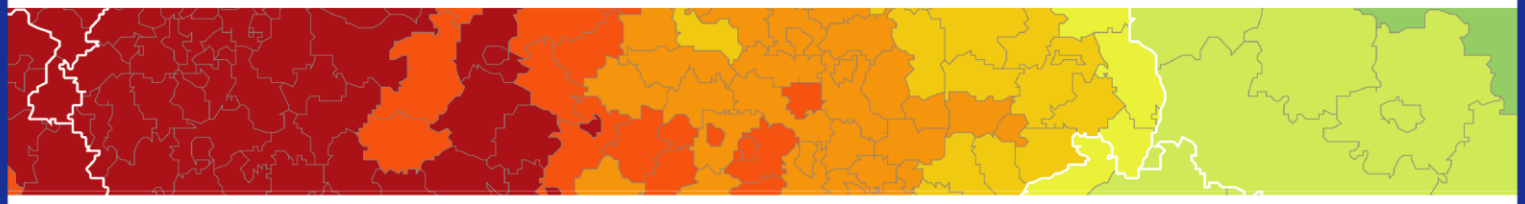
Based on the review on broader big data landscape and the case studies exploring the potentials related to three different datasets, the final report of this targeted analysis presents a summary of policy implications as well as recommendations for big data-driven corridor development. Figure 2 presents a summary of measures that can be promoted through corridor collaboration and territorial development. The first measure is to strategize for data-driven territorial governance, as well as to create actions plans how develop comprehensive data utilisation. The second step is to build capacity to improve administrators' capacity to recognize potentials in new data sources and generate actionable insights from them. In addition, public organisations should develop their skills in data analytics. The third measure is to establish and foster collaboration and partnerships. Partnerships can include public-private partnerships to incentivise data provision, as well as partnerships in research and development, - e.g., between research organisations and companies. The fourth step is to enable big data utilisation. This can be accomplished by harmonising and standardising data management systems, diminishing or clarifying legislative restrictions on wider data utilisation, as well as lobbying for long-term theoretical research on new data sources and methods.

The measures are further elaborated upon in the practical guide of this targeted analysis and address collective and individual actions needed to promote big data utilisation. Producing new evidence to support policy-making using big data is a collective effort that needs to be supported by the EU and national level authorities. At the end, however, the utilisation potential

is also relative to the intended use: public-private partnerships with companies may provide rapid access to data insights for policy-making, whereas harmonisation of statistical systems at the EU level supports integrated territorial development in the long run. In addition, caution needs to be paid to balancing between experimenting, standardizing and resourcing big data utilisation due to the rapid evolution of data economy and costs related to data utilisation.

Figure 2: Steps towards data-driven corridor governance.





ESPON 2020 – More information

ESPON EGTC
4 rue Erasme, L-1468 Luxembourg - Grand Duchy of Luxembourg
Phone: +352 20 600 280
Email: **info@espon.eu**
www.espon.eu, **Twitter**, **LinkedIn**, **YouTube**

The ESPON EGTC is the Single Beneficiary of the ESPON 2020 Cooperation Programme. The Single Operation within the programme is implemented by the ESPON EGTC and co-financed by the European Regional Development Fund, the EU Member States and the Partner States, Iceland, Liechtenstein, Norway and Switzerland.